

THE LARGE DEVIATION PRINCIPLE
FOR
MEASURES WITH RANDOM WEIGHTS

R. S. Ellis

*Department of Mathematics and Statistics, University of Massachusetts,
Amherst, Massachusetts 01003, U.S.A.*

J. Gough and J. V. Pulé [†]

*Department of Mathematical Physics, University College Dublin,
Belfield, Dublin 4, Ireland*

Abstract: In this paper we study the problem of large deviations for measures with random weights. We are motivated by previous work dealing with the special case occurring in the statistical mechanics of the Bose gas. We study the problem in an abstract setting, isolating what is general from what is dependent on Bose statistics. We succeed in proving the large deviation principle for a large class of measures with random weights and obtaining the corresponding rate function in an explicit form. In particular our results are applicable to the Fermi gas and the spherical model.

[†] *Research Associate, School of Theoretical Physics, Dublin Institute for Advanced Studies.*

§1. Introduction

Measures with random weights arise naturally in statistical mechanics. By *measures with random weights* we mean measures of the form

$$m = \sum_j X_j \delta_{a_j},$$

where the a_j 's are fixed points and the X_j 's are random variables. These are to be contrasted with *empirical measures* where the opposite is the case, the X_j 's being fixed numbers and the a_j 's random variables. While there is a large literature on large deviation results for empirical measures (see, for example, [2] and [3]), not many people have addressed the problem of large deviations for measures with random weights. This problem has been studied mainly in the context of the Bose gas [4, 5]. For the Bose gas the points a_j represent different momenta or energy levels, while the random variables X_j represent the number of particles at each a_j corresponding to Bose statistics. The present paper is motivated by the results of [4]. While we follow the general outline of [4], here we are interested in studying the problem in an abstract setting, isolating what is general from what is dependent on Bose statistics. We succeed in proving the large deviation principle for a large class of measures with random weights and obtaining the corresponding rate function in an explicit form. A benefit of our general approach is that the results of this paper also apply to the Fermi gas [6] and the spherical model. We shall describe these, together with the Bose gas, after we have set up the problem.

Let σ be a positive Borel measure on the closed halfline \mathbf{R}_+ and for $s \in \mathbf{R}$ we define

$$\pi(s) \doteq \ln \int_{\mathbf{R}_+} e^{sy} \sigma(dy). \quad (1.1)$$

Let $\gamma \doteq \sup\{s \in \mathbf{R} : \pi(s) < \infty\}$ and assume that $\gamma > -\infty$. The function π is lower semi-continuous, convex and on $(-\infty, \gamma)$ it is C^∞ . We shall assume that if $\gamma < \infty$, then $\lim_{s \uparrow \gamma} \pi'(s) = \infty$.

Let X be a locally compact Hausdorff space and let ξ be a function mapping X into \mathbf{R} . We assume that ξ satisfies the following conditions:

Hypothesis 1.

- (i) ξ is continuous.
- (ii) $-\xi$ has compact level sets; i.e. for each $b < \infty$, the set $\{x \in X : -\xi(x) \leq b\}$ is compact.
- (iii) $\xi_0 \doteq \sup_{x \in X} \xi(x) < \gamma$.

For each $n \in \mathbb{N}$ let $\{x_j(n) : j = 1, 2, \dots\}$ be a countable subset of X and let $a_n \in \mathbb{R}$ be such that $a_n \rightarrow \infty$ as $n \rightarrow \infty$. We assume that if A is a compact subset of X , then $|\{j \in \mathbb{N} : x_j(n) \in A\}|$ is finite, where $|\cdot|$ denotes cardinality. Define a measure μ_n on X by the formula

$$\mu_n(A) \doteq (a_n)^{-1} |\{j \in \mathbb{N} : x_j(n) \in A\}| \quad (1.2)$$

for every Borel subset A of X . Let \mathcal{G} be the family of continuous functions g mapping X into \mathbb{R} and satisfying

$$\sup_{x \in X} g(x) < \gamma$$

and

$$\sup_{x \in X} |g(x) - a\xi(x)| < b$$

for some positive real numbers a and b depending on g .

The following more or less standard definitions are used in this paper:

A positive Borel measure ν on X is a positive Radon measure if

- (i) $\nu(K) < \infty$ for every compact subset K of X .
- (ii) For every Borel set $A \subset X$

$$\nu(A) = \inf\{\nu(V) : A \subset V, V \text{ open}\}.$$

- (iii) For every Borel set $A \subset X$ such that A is open or $\nu(A) < \infty$

$$\nu(A) = \sup\{\nu(K) : K \subset A, K \text{ compact}\}.$$

A positive Radon measure ν is said to be regular if (iii) is satisfied for every Borel set $A \subset X$.

A measure ν is said to be a bounded Radon measure if it can be expressed in the form $\nu = \nu_1 - \nu_2$ where ν_1 and ν_2 are positive bounded Radon measures.

Let \tilde{E} be the space of bounded Radon measures on X . For $m \in \tilde{E}$ and $f \in \mathcal{C}^b(X)$, let

$$\langle m, f \rangle \doteq \int_X f(x) m(dx). \quad (1.3)$$

We can define a norm $\|\cdot\|$ on \tilde{E} by the formula

$$\|m\| \doteq \sup\{|\langle m, f \rangle| : f \in \mathcal{C}^b(X), \|f\|_\infty = 1\}.$$

Let E be the set of positive bounded Radon measures on X . We note that if $m \in E$, then m is regular and $\|m\| = m(X)$. Here we equip E with the narrow

topology. The narrow topology is the weakest topology for which the mappings $m \mapsto \langle m, f \rangle$ are continuous for all f in $\mathcal{C}^b(X)$.

In order to formulate our large deviation theorem, we shall also assume that there is a positive regular Radon measure μ on X satisfying the following conditions:

Hypothesis 2.

- (i) $\text{supp } \mu = X$.
- (ii) For each $g \in \mathcal{G}$,

$$\int_X |\pi(g(x))| \mu(dx) < \infty \quad \text{and} \quad \int_X |\xi(x)| \pi'(g(x)) \mu(dx) < \infty.$$

- (iii) For each $g \in \mathcal{G}$,

$$\lim_{n \rightarrow \infty} \int_X \pi(g(x)) \mu_n(dx) = \int_X \pi(g(x)) \mu(dx).$$

The following lemma gives some useful consequences of Hypothesis 2. Because the lemma follows fairly easily from the convexity of π , we do not give the proof here but save it for an appendix.

Lemma 1.1 *Suppose that Hypothesis 2 is satisfied. Then the following statements hold.*

- (i) For each $g \in \mathcal{G}$

$$\int_X \pi'(g(x)) \mu(dx) < \infty.$$

- (ii) For each $g \in \mathcal{G}$ and $f \in \mathcal{C}^b(X)$

$$\lim_{n \rightarrow \infty} \int_X f(x) \pi'(g(x)) \mu_n(dx) = \int_X f(x) \pi'(g(x)) \mu(dx).$$

- (iii) If $\{c_n\}$ is a sequence of real numbers converging to zero, then for each $g \in \mathcal{G}$ and $f \in \mathcal{C}^b(X)$

$$\lim_{n \rightarrow \infty} c_n^{-1} \int_X \{\pi(g(x) + c_n f(x)) - \pi(g(x))\} \mu_n(dx) = \int_X f(x) \pi'(g(x)) \mu(dx).$$

For each $n \in \mathbb{N}$ let $\{X_j^{(n)} : j = 1, 2, \dots\}$ be positive independent random variables, $X_j^{(n)}$ having distribution $\sigma_j^{(n)}$, where

$$\sigma_j^{(n)}(dy) \doteq \frac{e^{\xi(x_j^{(n)})y} \sigma(dy)}{\int_{\mathbb{R}_+} e^{\xi(x_j^{(n)})y} \sigma(dy)}. \quad (1.4)$$

Let \mathbf{P}_n be the corresponding product measure on $\Omega \doteq \mathbf{R}_+^{\mathbf{N}}$ and let $\tilde{\Omega} \doteq \{\omega \in \Omega : \sum_{j \geq 1} X_j^{(n)}(\omega) < \infty\}$. Since

$$\mathbb{E}(\sum_{j \geq 1} X_j^{(n)}) = a_n \int_X \pi'(\xi(x)) \mu_n(dx) \quad (1.5)$$

and the integral in (1.5) is finite by conditions (ii) and (iii) in Hypothesis 2, we have $\mathbf{P}_n(\tilde{\Omega}) = 1$. For each $\omega \in \tilde{\Omega}$, define the bounded measure $L_n(\omega, \cdot)$ on X by the formula

$$L_n(\omega, A) \doteq a_n^{-1} \sum_{j \geq 1} X_j^{(n)}(\omega) \delta_{x_j(n)}(A) \quad (1.6)$$

for every Borel subset $A \subset X$. L_n takes values in E . Finally let \mathbf{K}_n be the probability measure induced by L_n on E ; i.e.,

$$\mathbf{K}_n \doteq \mathbf{P}_n \circ L_n^{-1}. \quad (1.7)$$

One of our goals is to prove that the sequence of probability measures $\{\mathbf{K}_n\}$ on E satisfies the large deviation principle. Before formulating this, we will specify, in three important examples that arise in statistical mechanics, the quantities σ , π , ξ and $x_j(n)$ appearing in the general definitions.

The Bose Gas

For the Bose gas σ is the counting measure, $\sigma(A) \doteq |A \cap \mathbf{N}|$. Hence for $s < 0$

$$\pi(s) = \ln \left(\sum_{j=0}^{\infty} e^{js} \right) = -\ln(1 - e^s).$$

Thus $\gamma = 0$. We also set $X \doteq \mathbf{R}^d$ and $\xi(x) \doteq \alpha - \|x\|^2$ for some $\alpha < 0$. Hypothesis 1 is satisfied. The set $\{x_j(n)\}$ is $\{2\pi n^{-\frac{1}{d}} k : k \in \mathbf{Z}^d\}$, so that if $a_n = n$, then μ_n converges in the sense of Condition (iii) of Hypothesis 2 to $\mu \doteq (2\pi)^{-d} m$, where m is Lebesgue measure on \mathbf{R}^d . Hypothesis 2 is satisfied. In this model the measures in E are interpreted as the occupation densities for the momentum states corresponding to $\{n^{-\frac{1}{d}} k : k \in \mathbf{Z}^d\}$.

An important objective in statistical mechanics is to obtain the grand canonical pressure p in the thermodynamic limit. For some Bose models, p can be expressed in the form

$$p \doteq \lim_{n \rightarrow \infty} \frac{1}{a_n} \ln \int_E e^{a_n G(m)} \mathbf{K}_n(dm),$$

where G is given by the formula

$$G(m) \doteq \zeta \|m\| - \frac{1}{2} \int_X \int_X v(x, x') m(dx) m(dx'). \quad (1.7a)$$

Here $\zeta \in \mathbf{R}$ and v is a bounded, continuous, positive definite function mapping X^2 into \mathbf{R} . If the topology on E is chosen so that G is continuous, then one can use Varadhan's Theorem [1, 2] to obtain a variational expression for p . A suitable topology is the narrow topology.

The Fermi gas

For the Fermi gas $\sigma(A) \doteq |A \cap \{0, 1\}|$,

$$\pi(s) \doteq \ln(1 + e^s)$$

(so that $\gamma = \infty$) and $\xi(x) \doteq \alpha - \|x\|^2$ for some $\alpha \in \mathbf{R}$. The other quantities are the same as for the Bose gas. Hypotheses 1 and 2 satisfied.

The Spherical Model

In this model

$$\sigma(dy) \doteq \frac{1}{\sqrt{y\pi}} dy.$$

Hence for $s < 0$

$$\pi(s) = \ln \int_{\mathbf{R}_+} e^{sy} \frac{1}{\sqrt{y\pi}} dy = -\frac{1}{2} \ln(-s).$$

Thus $\gamma = 0$. Let $\{c_1, c_2, \dots, c_d\}$ be a basis for \mathbf{R}^d and let Λ be the Bravais lattice generated by this basis:

$$\Lambda \doteq \left\{ \sum_{i=1}^d m_i c_i : m \in \mathbf{Z}^d \right\}.$$

For $n \in \mathbf{N}$ let Λ_n be the subset of Λ given by

$$\Lambda_n \doteq \left\{ \sum_{i=1}^d m_i c_i : m \in \{-n, -n+1, \dots, n-1, n\}^d \right\}.$$

We define a_n to be the number of lattice points in Λ_n ; that is $a_n = (2n+1)^d$. We also choose a positive function $u : \Lambda \mapsto \mathbf{R}$ such that $\sum_{y \in \Lambda} u(y) < \infty$. Define $\{b_1, b_2, \dots, b_d\}$ to be the basis of \mathbf{R}^d satisfying $\langle c_i, b_j \rangle = 2\pi \delta_{ij}$, and let Λ^r be the parallelepiped

$$\Lambda^r \doteq \left\{ \sum_{i=1}^d x_i b_i : x \in \mathbf{R}^d, |x_i| \leq \frac{1}{2}, i = 1, \dots, d \right\}.$$

Define a function $\tilde{\xi} : \Lambda^r \mapsto \mathbf{R}$ by the formula

$$\tilde{\xi}(x) \doteq \sum_{y \in \Lambda} u(y) \sin^2 \frac{1}{2} \langle x, y \rangle.$$

We set $X \doteq \Lambda^r$ and $\xi(x) \doteq \alpha - \tilde{\xi}(x)$ for some $\alpha < 0$. Hypothesis 1 is satisfied. Let Λ_n^r be the lattice reciprocal to Λ_n :

$$\Lambda_n^r = \left\{ (2n+1)^{-1} \sum_{j=1}^d m_j b_j : m \in \{-n, -n+1, \dots, n-1, n\}^d \right\}.$$

The set $\{x_j(n)\}$ then is equal to Λ_n^r , so that μ_n converges in the sense of condition (iii) of Hypothesis 2 to $\mu \doteq C^{-1}m$, where m is Lebesgue measure on Λ^r and C is the volume of Λ^r . Hypothesis 2 is satisfied. This completes our presentation of examples.

We return to the general development, recalling the probability measures \mathbf{K}_n on E defined in equation (1.7). The first objective of this paper is to prove that if E is equipped with the narrow topology, then the sequence of probability measures $\{\mathbf{K}_n\}$ on E obeys the large deviation principle [1, 2]. We recall that the sequence of probability measures $\{\mathbf{K}_n\}$ on E is said to obey the large deviation principle with constants $\{a_n\}$ and rate function $I : E \mapsto [0, \infty]$ if the following conditions are satisfied:

- (LD1) I is lower semi-continuous;
- (LD2) For each $b < \infty$, the level set $\{m \in E : I(m) \leq b\}$ is compact;
- (LD3) For each closed set C

$$\limsup_{l \rightarrow \infty} a_n^{-1} \ln \mathbf{K}_n(C) \leq -I(C);$$

- (LD4) For each open set G ,

$$\liminf_{l \rightarrow \infty} a_n^{-1} \ln \mathbf{K}_n(G) \geq -I(G).$$

Here we have used the notation

$$I(A) \doteq \inf_{m \in A} I(m) \tag{1.8}$$

for a non-empty subset A of E and we set $I(\emptyset) \doteq \infty$.

Let us go back to the example of the Bose gas discussed earlier and assume that the function G on E given in equation (1.7a) is continuous in the narrow topology

on E . Then the large deviation principle for $\{\mathbf{K}_n\}$ and Varadhan's Theorem give a variational formula for the pressure p ; namely

$$p \doteq \lim_{n \rightarrow \infty} \frac{1}{a_n} \ln \int_E e^{a_n G(m)} \mathbf{K}_n(dm) = \sup_{m \in E} [G(m) - I(m)].$$

Knowledge of the minimizers of this variational expression can give great insight into the physical properties of the equilibrium states of the model. Clearly, an explicit form for the rate function I is very helpful in the study of the variational problem. The second objective of this paper is thus to obtain an explicit formula for the rate function in the general case.

This paper is set out as follows. In Section 2 we shall prove that $\{\mathbf{K}_n\}$ satisfies the large deviation principle (Theorem 2) and give an explicit formula for the rate function I . (Theorem 3). The proofs of Theorems 2 and 3 depend crucially on the Approximation Theorem, stated in Theorem 1 and proved in Section 3. Lemma 1.1 is proved in an appendix.

§2. Large Deviations

In this section we prove the large deviation principle for the sequence of measures $\{\mathbf{K}_n\}$ and obtain an explicit form for the rate function I . These results are based on the Approximation Theorem which is stated in Theorem 1 in this section and proved in Section 3. The large deviation principle is stated in Theorem 2.

Let $\mathcal{D} \doteq \{f : f \in C^b(X), \sup(\xi(x) + f(x)) < \gamma\}$. For $f \in \mathcal{D}$, define

$$C_n(f) \doteq \frac{1}{a_n} \ln \int_E e^{a_n \langle m, f \rangle} \mathbf{K}_n(dm) = \frac{1}{a_n} \ln E\left\{\exp\left(\sum_{j \geq 1} f(x_j(n)) X_j^n\right)\right\}. \quad (2.1)$$

Then

$$C_n(f) = \int_X \{\pi(\xi(x) + f(x)) - \pi(\xi(x))\} \mu_n(dx). \quad (2.2)$$

If we define

$$C(f) \doteq \lim_{n \rightarrow \infty} C_n(f), \quad (2.3)$$

then by condition (iii) of Hypothesis 2

$$C(f) = \int_X \{\pi(\xi(x) + f(x)) - \pi(\xi(x))\} \mu(dx). \quad (2.4)$$

Note that by using the convexity of π and Lemma 1.1 (i), one may easily check that C is continuous with respect to the supremum norm $\|\cdot\|_\infty$ on \mathcal{D} . For $m \in E$ we define

$$I(m) \doteq \sup_{f \in \mathcal{D}} \{\langle m, f \rangle - C(f)\}. \quad (2.5)$$

In Theorem 2, we shall prove that I is the rate function in the large deviation principle for $\{\mathbf{K}_n\}$.

For $f \in \mathcal{D}$ define

$$\rho^f(x) \doteq \pi'(\xi(x) + f(x)) \quad (2.6)$$

and

$$m^f(dx) \doteq \pi'(\xi(x) + f(x)) \mu(dx); \quad (2.7)$$

i.e. m^f is the element of E which is absolutely continuous with respect to μ and has density ρ^f . In Section 3 we shall prove the next theorem.

Theorem 1. (Approximation Theorem) *Let μ be a positive regular Radon measure on X satisfying Hypothesis 2. Let m be an element of E such that $I(m)$ is finite. The following conclusions hold.*

(a) If m is absolutely continuous with respect to μ and has density ρ , then there is a sequence $\{f_n\}$ in \mathcal{D} such that

1. $\lim_{n \rightarrow \infty} \int_X |\rho(x) - \rho^{f_n}(x)| \mu(dx) = 0$,
2. $\lim_{n \rightarrow \infty} I(m^{f_n}) = I(m)$.

(b) If N is a neighbourhood of m and $\epsilon > 0$, then there exists $f \in \mathcal{D}$ such that

1. $m^f \in N$,
2. $|I(m^f) - I(m)| < \epsilon$.

In order to prove the large deviation principle, we shall also need the following four lemmas. Choose a number $\tilde{\gamma} \in (\xi_0, \gamma)$ and for each $k \in \mathbb{N}$ define the function $f_k : X \mapsto \mathbb{K}$ by the formula

$$f_k(x) \doteq \begin{cases} 0 & \text{if } \xi(x) \geq -k + \tilde{\gamma}, \\ -k(\xi(x) + k - \tilde{\gamma})/2 & \text{if } -k + \tilde{\gamma} - 1 < \xi(x) < -k + \tilde{\gamma}, \\ k/2 & \text{if } \xi(x) \leq -k + \tilde{\gamma} - 1. \end{cases} \quad (2.8)$$

We note that

$$0 \leq f_k(x) \leq \frac{1}{2}k.$$

If $\xi(x) < -k + \tilde{\gamma}$, then

$$f_k(x) \leq \frac{1}{2}k < \frac{1}{2}(\tilde{\gamma} - \xi(x)).$$

Since $\frac{1}{2}(\tilde{\gamma} - \xi(x)) > 0$, when $\xi(x) \geq -k + \tilde{\gamma}$, we have $f_k(x) = 0 < \frac{1}{2}(\tilde{\gamma} - \xi(x))$. Therefore for all $x \in X$

$$f_k(x) + \xi(x) \leq \frac{1}{2}(\xi(x) + \tilde{\gamma}) \leq \frac{1}{2}(\xi_0 + \tilde{\gamma}) < \gamma,$$

and so $f_k \in \mathcal{D}$. Let $f_{\max}(x) \doteq \sup_{k \geq 1} f_k(x)$. Then

$$0 \leq f_{\max}(x) \leq \frac{1}{2}(\tilde{\gamma} - \xi(x)).$$

Note that f_{\max} need not be in \mathcal{D} . For $M > 0$ define

$$\Phi_M \doteq \bigcap_{k \geq 1} \{m \in F : \langle m, f_k \rangle \leq M\} = \{m \in E : \sup_{k \geq 1} \langle m, f_k \rangle \leq M\}$$

and

$$B_M \doteq \{m \in E : \|m\| \leq M\},$$

and put $W_M \doteq \Phi_M \cap B_M$. Φ_M is clearly closed. We shall prove that W_M is compact after the next lemma. We need the following definition.

Definition: A set $\Phi \subset E$ is *uniformly tight* if given $\epsilon > 0$, there is a compact subset $\Gamma \subset X$ such that $m(\Gamma^c) < \epsilon$ for all $m \in \Phi$.

Lemma 2.1 *The set Φ_M is uniformly tight.*

Proof: Let $m \in \Phi_M$. Then $\langle m, f_k \rangle \leq M$ for all $k \in \mathbb{N}$, and so

$$\frac{1}{2}k \int_{\{x: \xi(x) < -(k+1-\tilde{\gamma})\}} m(dx) = \int_{\{x: \xi(x) < -(k+1-\tilde{\gamma})\}} f_k(x) m(dx) \leq \langle m, f_k \rangle \leq M$$

for all $k \in \mathbb{N}$. Hence given $\epsilon > 0$ there exists $n \in \mathbb{N}$ such that $m\{x : \xi(x) < -(n+1-\tilde{\gamma})\} < \epsilon$ for all m in Φ_M ; but $\{x : \xi(x) \geq -(n+1-\tilde{\gamma})\}$ is compact by condition (ii) of Hypothesis 1. This completes the proof of the lemma with

$$\Gamma \doteq \{x \in X : \xi(x) \geq -(n+1-\tilde{\gamma})\}.$$

□

We now prove that W_M is compact.

Lemma 2.2 *The set W_M is compact in the narrow topology.*

Proof: Since $W_M \subset B_M$, the set W_M is uniformly bounded; since $W_M \subset \Phi_M$, the set W_M is uniformly tight. Since X is a locally compact Hausdorff space, it follows from Prokhorov's Criterion (Theorem 1 of Number 5.5 of [7]) that W_M is compact.

□

In order to prove the upper large deviation bound, we want to show that the sequence of measures $\{\mathbb{K}_n\}$ is exponentially tight; that is, given $L \in (0, \infty)$ there exists a compact subset A_L of X such that

$$\limsup_{n \rightarrow \infty} \frac{1}{a_n} \ln \mathbb{K}_n(A_L^c) \leq -L.$$

This is carried out in the next lemma.

Lemma 2.3 *The sequences of measures $\{\mathbb{K}_n\}$ is exponentially tight.*

Proof: For each $M \in (0, \infty)$ the set $W_M = B_M \cap \Phi_M$ is compact and $W_M^c = B_M^c \cup \Phi_M^c$. Hence it suffices to prove that

$$\lim_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{a_n} \ln \mathbf{K}_n(B_M^c) = -\infty \quad (2.9)$$

and

$$\lim_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{a_n} \ln \mathbf{K}_n(\Phi_M^c) = -\infty. \quad (2.10)$$

Choosing $\alpha \in (0, \gamma - \xi_0)$ we have

$$\begin{aligned} \mathbf{K}_n(B_M^c) &= \int_{\{m \in E: \|m\| > M\}} \mathbf{K}_n(dm) \leq \int_{\{m \in E: \|m\| > M\}} e^{a_n \alpha (\|m\| - M)} \mathbf{K}_n(dm) \\ &\leq \int e^{a_n \alpha (\|m\| - M)} \mathbf{K}_n(dm) = e^{-a_n \alpha M} e^{a_n C_n(\alpha)}. \end{aligned}$$

Therefore

$$\limsup_{n \rightarrow \infty} \frac{1}{a_n} \ln \mathbf{K}_n(B_M^c) \leq -\alpha M + C(\alpha)$$

and since $\alpha > 0$, the limit (2.9) follows.

We have

$$\begin{aligned} \frac{1}{a_n} \int_E e^{a_n \sup_{k \geq 1} \langle m, f_k \rangle} \mathbf{K}_n(dm) &\leq \frac{1}{a_n} \ln \int_E e^{a_n \int_X f_{\max}(x) m(dx)} \mathbf{K}_n(dm) \\ &\leq \frac{1}{a_n} \ln \int_E e^{a_n \int_X \frac{1}{2}(\tilde{\gamma} - \xi(x)) m(dx)} \mathbf{K}_n(dm) \\ &= C_n \left(\frac{1}{2}(\tilde{\gamma} - \xi(x)) \right) \\ &= \int_X \left\{ \pi \left(\frac{1}{2} \xi(x) + \frac{1}{2} \tilde{\gamma} \right) - \pi(\xi(x)) \right\} \mu_n(dx). \end{aligned}$$

Hence

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{a_n} \ln \int_E e^{a_n \sup_{k \geq 1} \langle m, f_k \rangle} \mathbf{K}_n(dm) \\ \leq \int_X \left\{ \pi \left(\frac{1}{2} \xi(x) + \frac{1}{2} \tilde{\gamma} \right) - \pi(\xi(x)) \right\} \mu(dx) \equiv A < \infty. \end{aligned} \quad (2.11)$$

Since by Chebyshev's Inequality

$$\mathbf{K}_n(\Phi_M^c) \leq e^{-a_n M} \int_E e^{a_n \sup_{k \geq 1} \langle m, f_k \rangle} \mathbf{K}_n(dm),$$

we have

$$\limsup_{n \rightarrow \infty} \frac{1}{a_n} \ln \mathbf{K}_n(\Phi_M^c) \leq A - M$$

and the limit (2.10) follows. This completes the proof of the exponential tightness of the sequence $\{\mathbb{K}_n\}$. □

The next lemma is needed in the proof of the large deviation lower bound. We shall use the following notation. For $m \in E$ and $f \in \mathcal{D}$ let

$$I(m, f) \doteq \langle m, f \rangle - C(f).$$

Lemma 2.4 For $f \in \mathcal{D}$

$$I(m^f) = I(m^f, f);$$

that is, for $m = m^f$ the supremum in the definition of $I(m^f)$ is attained at f .

Proof. For fixed $x \in X$, define $g : \mathbb{R} \mapsto \mathbb{R}$ by the formula

$$g(y) \doteq \pi'(\xi(x) + f(x))y - \pi(\xi(x) + y).$$

Then

$$g'(y) = \pi'(\xi(x) + f(x)) - \pi'(\xi(x) + y)$$

and thus $g'(f(x)) = 0$. But g is concave and therefore $g(y) \leq g(f(x))$. Letting $y = h(x)$, where $h \in \mathcal{D}$, and integrating with respect to μ , we get

$$\int_X g(h(x))\mu(dx) \leq \int_X g(f(x))\mu(dx).$$

This equivalent to

$$\int_X \pi'(\xi(x) + f(x))h(x)\mu(dx) - C(h) \leq \int_X \pi'(\xi(x) + f(x))f(x)\mu(dx) - C(f)$$

or

$$I(m^f, h) \leq I(m^f, f).$$

It follows that $I(m^f) = I(m^f, f)$, as claimed. □

We are now ready to prove the large deviation principle.

Theorem 2. The sequence of probability measures $\{\mathbb{K}_n\}$ on E satisfies the large deviation principle with constants $\{a_n\}$ and rate function

$$I(m) = \sup_{f \in \mathcal{D}} \{\langle m, f \rangle - C(f)\}.$$

Proof: We first verify (LD1)-(LD2). I is lower semi-continuous because the supremum of a family of continuous functions is lower semi-continuous. Hence (LD1) holds. To prove (LD2) (compact level sets of I), we first note that the lower semi-continuity of I implies that the level set $S_b \doteq \{m \in E : I(m) \leq b\}$ is closed. For $f \in \mathcal{D}$ and m in S_b we have

$$b \geq I(m) \geq \langle m, f \rangle - C(f). \quad (2.12)$$

Choosing $\alpha \in (0, \gamma - \xi_0)$ and putting $f(x) = \alpha$, we have

$$b \geq \alpha \|m\| - \int_X (\pi(\xi(x) + \alpha) - \pi(\xi(x))) \mu(dx).$$

It follows that if M is chosen large enough so that

$$M > \alpha^{-1} \left(b + \int_X \{\pi(\xi(x) + \alpha) - \pi(\xi(x))\} \mu(dx) \right),$$

then $S_b \subset B_M$. Also putting $f \doteq f_k$ in (2.12) (the function f_k is defined in equation (2.8)), we obtain

$$b \geq \langle m, f_k \rangle - C(f_k).$$

But by definition of the constant A (see equation (2.11))

$$\begin{aligned} C(f_k) &= \int_X \{\pi(\xi(x) + f_k(x)) - \pi(\xi(x))\} \mu(dx) \\ &\leq \int_X \left\{ \pi \left(\frac{1}{2} \xi(x) + \frac{1}{2} \tilde{\gamma} \right) - \pi(\xi(x)) \right\} \mu(dx) = A, \end{aligned}$$

Hence $\langle m, f_k \rangle \leq b + A$ for all $k \geq 1$. It follows that $S_b \subset \Phi_M$ for all $M \geq A + b$. We have thus proved that, for M sufficiently large, the level set S_b is a proper subset of W_M and that S_b is closed. It follows from Lemma 2.2 that S_b is compact. Hence (LD2) holds.

We now prove (LD3), the large deviation upper bound for closed sets. Lemma 2.3 proved that the sequence of measures $\{\mathbf{K}_n\}$ is exponentially tight. Hence by Lemma 2.1.5 in Deuschel-Stroock [3], it suffices to prove the large deviation upper bound for compact subsets of E . In order to carry this out, we follow Lemma VII.4.1 in [2] and make use of the next lemma, whose proof is essentially identical and therefore omitted. We merely remark that this lemma uses the continuity, with respect to the supremum norm $\|\cdot\|_\infty$, of C at 0.

Lemma 2.5 Given $f \in \mathcal{D}$ and $\beta \in \mathbb{R}$, define $H_+(f, \beta)$ by the formula

$$H_+(f, \beta) \doteq \{m \in E : \langle m, f \rangle - C(f) \geq \beta\}.$$

Let K be a compact subset of E . Then for any number $\beta < I(K)$ there exists a finite set f_1, \dots, f_r of non-zero elements of \mathcal{D} such that

$$K \subset \bigcup_{j=1}^r H_+(f_j, \beta).$$

The rest of the proof of (LD3) is standard; we give it for the sake of completeness. Let K be a compact subset of E . For each $\beta < I(K)$ we have by Lemma 2.5 and Chebyshev's Inequality

$$\begin{aligned} \mathbb{K}_n(K) &\leq \sum_{j=1}^r \mathbb{K}_n(H_+(f_j, \beta)) \\ &\leq \sum_{j=1}^r e^{-a_n(C(f_j) + \beta)} \int_E e^{a_n \langle m, f_j \rangle} \mathbb{K}_n(dm) \\ &= e^{-a_n \beta} \sum_{j=1}^r e^{a_n [C_n(f_j) - C(f_j)]}. \end{aligned}$$

Hence

$$\limsup_{n \rightarrow \infty} \frac{1}{a_n} \ln \mathbb{K}_n(K) \leq -\beta.$$

Since this holds for all $\beta < I(K)$, we have

$$\limsup_{n \rightarrow \infty} \frac{1}{a_n} \ln \mathbb{K}_n(K) \leq -I(K).$$

This completes the proof of the large deviation upper bound for the compact set K . Thus (LD3) holds.

We now prove (LD4), the large deviation lower bound for open sets. Let G be an open subset of E . If $I(G) = \infty$, then (LD4) holds for G . So we suppose that $I(G) < \infty$. Then for each $\epsilon > 0$ there exists a measure $m \in G$ such that $I(m) \leq I(G) + \epsilon$. By Theorem 1(b) there exists $f \in \mathcal{D}$ such that $m^f \in G$ and $I(m^f) < I(m) + \epsilon$, so that

$$I(m^f) < I(G) + 2\epsilon. \tag{2.13}$$

Now let

$$G_\epsilon \doteq G \cap \{m \in E : |\langle m, f \rangle - \langle m^f, f \rangle| < \epsilon\}$$

and for $n \in \mathbb{N}$ define the measures \tilde{K}_n on E by the formula

$$\tilde{K}_n(dm) \doteq e^{a_n\{\langle f, m \rangle - C_n(f)\}} K_n(dm).$$

Recalling that

$$e^{a_n C_n(f)} = \int_E e^{a_n \langle m, f \rangle} K_n(dm),$$

we see that \tilde{K}_n is a probability measure on E . We shall prove that for all n sufficiently large $\tilde{K}_n(G_\epsilon) > \frac{1}{2}$. Since G_ϵ is open and $m^f \in G_\epsilon$, there exist $f_1, \dots, f_r \in C^b(X)$ and $\delta > 0$ such that

$$N_\delta \doteq \bigcap_{j=1}^r \{m \in E : |\langle f_j, m - m^f \rangle| < \delta\} \subset G_\epsilon.$$

Define the function $g : E \mapsto \mathbb{R}^r$ by the formula

$$g(m) \doteq ((g(m))_1, (g(m))_2, \dots, (g(m))_r),$$

where for each $j \in \{1, 2, \dots, r\}$

$$(g(m))_j \doteq \langle f_j, m - m^f \rangle.$$

Then define $Q_n \doteq \tilde{K}_n \circ g^{-1}$. For real numbers s_1, s_2, \dots, s_r , the Laplace transform of the probability measure Q_n is defined by the formula

$$\begin{aligned} \hat{Q}_n(s_1, s_2, \dots, s_r) &\doteq \int_{\mathbb{R}^r} e^{-\sum_{j=1}^r s_j t_j} Q_n(dt) \\ &= \int_E e^{-\langle \sum_{j=1}^r s_j f_j, m - m^f \rangle} \tilde{K}_n(dm) \\ &= e^{\langle \sum_{j=1}^r s_j f_j, m^f \rangle} \exp \left[a_n \left\{ C_n(f - \frac{1}{a_n} \sum_{j=1}^r s_j f_j) - C_n(f) \right\} \right]. \end{aligned}$$

By Lemma 1.1 (iii)

$$\begin{aligned} \lim_{n \rightarrow \infty} a_n \left\{ C_n \left(f - \frac{1}{a_n} \sum_{j=1}^r s_j f_j \right) - C_n(f) \right\} &= - \int_X \sum_{j=1}^r s_j f_j(x) \pi'(\xi(x) + f(x)) \mu(dx) \\ &= - \langle \sum_{j=1}^r s_j f_j, m^f \rangle. \end{aligned}$$

Hence for $(s_1, s_2, \dots, s_r) \in \mathbb{R}^r$

$$\lim_{n \rightarrow \infty} \hat{Q}_n(s_1, s_2, \dots, s_r) = 1$$

and so by Chebyshev's Inequality

$$\lim_{n \rightarrow \infty} Q_n\{\mathbf{R}^r \setminus [-\delta, \delta]^r\} = 0.$$

Since

$$\tilde{K}_n(G_\epsilon) \geq \tilde{K}_n(N_\delta) = \tilde{K}_n \circ g^{-1}([-\delta, \delta]^r) = Q_n([-\delta, \delta]^r)$$

and $Q_n([-\delta, \delta]^r) \rightarrow 1$ as $n \rightarrow \infty$, we have that $\tilde{K}_n(G_\epsilon) > \frac{1}{2}$ for all n sufficiently large. Now

$$K_n(G) \geq K_n(G_\epsilon) = e^{a_n C_n(f)} \int_{G_\epsilon} e^{-a_n \langle m, f \rangle} \tilde{K}_n(dm) \geq e^{a_n [C_n(f) - \langle m^f, f \rangle - \epsilon]} \tilde{K}_n(G_\epsilon).$$

Therefore by Lemma 2.4, we get

$$\liminf_{n \rightarrow \infty} \frac{1}{a_n} \ln K_n(G) \geq C(f) - \langle m^f, f \rangle - \epsilon = -I(m^f) - \epsilon \geq -I(G) - 3\epsilon.$$

The last inequality follows from (2.13). Since ϵ is arbitrary, (LD4) holds. We have completed the proof of the large deviation principle for the measures $\{K_n\}$ with rate function I . □

As we remarked in Section 1, one application of large deviations in statistical mechanics is the use of Varadhan's Theorem to obtain a variational formula for the grand canonical pressure or the canonical free energy density. The variational problem is not studied here, but clearly it is very desirable for applications to have an explicit form for the rate function. The rest of this section is devoted to obtaining an explicit form for the rate function I , analogous that found in [4] for the Bose gas.

We split off that part of the measure m which is singular with respect to μ and deal with it separately. For $m \in E$, let $m = m_s + m_a$ be the Lebesgue decomposition of m with respect to μ into the singular part m_s and the absolutely continuous part m_a ; let ρ be the density of m_a so that $m_a(dx) = \rho(x)\mu(dx)$. Define $U : E \mapsto [0, \infty]$ by setting $U(0) = 0$ and for $m \neq 0$

$$U(m) \doteq \begin{cases} \int_X (\gamma - \xi(x)) m(dx) & \text{if } \gamma < \infty, \\ \infty & \text{if } \gamma = \infty. \end{cases}$$

We recall that $\gamma \doteq \sup\{s \in \mathbf{R} : \pi(s) < \infty\}$ and that by condition (iii) of Hypothesis 1, $\sup_{x \in X} \xi(x) < \gamma$.

The next lemma gives a useful formula for $U(m)$.

Lemma 2.6 For all $m \in E$, $\sup_{f \in \mathcal{D}} \langle m, f \rangle = U(m)$.

Proof: Clearly if $m = 0$, $\sup_{f \in \mathcal{D}} \langle m, f \rangle = 0$. Take $m \in E$, $m \neq 0$. Suppose $\gamma = \infty$ and let $f(x) = c$ for all $x \in X$. Then $f \in \mathcal{D}$ and $\langle m, f \rangle = c\|m\|$; thus $\sup_{f \in \mathcal{D}} \langle m, f \rangle \geq c\|m\|$ and since c is an arbitrary real number $\sup_{f \in \mathcal{D}} \langle m, f \rangle = \infty$. Now suppose $\gamma < \infty$. Since $f \in \mathcal{D}$, $\sup_{x \in X} (f(x) + \xi(x)) < \gamma$ and so

$$\sup_{f \in \mathcal{D}} \langle m, f \rangle \leq \int_X (\gamma - \xi(x)) m(dx). \quad (2.14)$$

In order to complete the proof, we show the opposite inequality. Let $\delta \in (0, \gamma - \delta_0)$ and for $n \in \mathbb{N} \cap ((\gamma - \delta_0)^{-1}, \infty)$ define

$$f_n(x) \doteq (\gamma - \xi(x) - n^{-1}) \wedge n$$

for all $x \in X$. Then $f_n \in \mathcal{D}$ and $f_n \geq 0$ and by Lebesgue's Monotone Convergence Theorem

$$\lim_{n \rightarrow \infty} \langle m, f_n \rangle = \int_X (\gamma - \xi(x)) m(dx).$$

Hence

$$\sup_{f \in \mathcal{D}} \langle m, f \rangle \geq \int_X (\gamma - \xi(x)) m(dx).$$

The proof of the lemma is complete. □

Lemma 2.7 For each $m \in E$,

$$I(m) = U(m_s) + I(m_a).$$

Proof: If $m_s = 0$ there is nothing to prove, we may therefore assume that $m_s \neq 0$. For all $f \in \mathcal{D}$,

$$\begin{aligned} I(m, f) &= \langle m, f \rangle - C(f) = \langle m_s, f \rangle + \langle m_a, f \rangle - C(f) \\ &= \langle m_s, f \rangle + I(m_a, f) \leq U(m_s) + I(m_a, f). \end{aligned}$$

Thus $I(m) \leq U(m_s) + I(m_a)$.

In order to complete the proof, we show the opposite inequality. Let $r < U(m_s)$ and choose $g \in \mathcal{D}$ such that $\langle g, m_s \rangle > r$. Let B be a subset of X such that $m_s(B^c) = 0$ and $\mu(B) = 0$ so that m_s is concentrated on B . For $n \in \mathbb{N}$ choose compact subsets $K_n \subset B$ such that $m_s(B \setminus K_n) < \frac{1}{n}$ and choose open subsets O_n , such that $B \subset O_{n+1} \subset O_n$, $m_a(O_n) < \frac{1}{n}$ and $\mu(O_n) < \frac{1}{n}$. We recall that this is

possible because m_a and m_s are bounded Radon measures and therefore regular. By Urysohn's lemma there exists a function $\tau_n \in C^b(X)$ such that $0 \leq \tau_n(x) \leq 1$ for all $x \in X$, $\tau_n(x) = 1$ for $x \in K_n$ and $\tau_n(x) = 0$ for $x \in O_n^c$. Let $f \in \mathcal{D}$ and define $f_n \in \mathcal{D}$ by

$$f_n(x) \doteq \tau_n(x)g(x) + (1 - \tau_n(x))f(x).$$

We then have

$$\begin{aligned} \langle m, f_n \rangle &= \langle m_s, g \rangle + \langle m_a, f \rangle + \int_X \{1 - \tau_n(x)\} \{f(x) - g(x)\} m_s(dx) \\ &\quad + \int_X \tau_n(x) \{g(x) - f(x)\} m_a(dx) \end{aligned}$$

Since $1 - \tau_n(x) = 0$ for $x \in K_n$ and $m_s(B^c) = 0$,

$$\begin{aligned} \int_X \{1 - \tau_n(x)\} \{f(x) - g(x)\} m_s(dx) &= \int_{X \setminus K_n} \{1 - \tau_n(x)\} \{f(x) - g(x)\} m_s(dx) \\ &= \int_{B \setminus K_n} \{1 - \tau_n(x)\} \{f(x) - g(x)\} m_s(dx). \end{aligned}$$

Similarly since $\tau_n(x) = 0$ for $x \in O_n^c$ and $m_a(B) = 0$,

$$\begin{aligned} \int_X \tau_n(x) \{g(x) - f(x)\} m_a(dx) &= \int_{O_n} \tau_n(x) \{g(x) - f(x)\} m_a(dx) \\ &= \int_{O_n \setminus B} \tau_n(x) \{g(x) - f(x)\} m_a(dx). \end{aligned}$$

Thus

$$\begin{aligned} \langle m, f_n \rangle &= \langle m_s, g \rangle + \langle m_a, f \rangle + \int_{B \setminus K_n} \{1 - \tau_n(x)\} \{f(x) - g(x)\} m_s(dx) \\ &\quad + \int_{O_n \setminus B} \tau_n(x) \{g(x) - f(x)\} m_a(dx) \\ &\geq \langle m_s, g \rangle + \langle m_a, f \rangle - 2\{\|f\|_\infty \vee \|g\|_\infty\} m_s(B \setminus K_n) \\ &\quad - 2\{\|f\|_\infty \vee \|g\|_\infty\} m_a(O_n) \\ &> r + \langle m_a, f \rangle - \frac{4}{n} \{\|f\|_\infty \vee \|g\|_\infty\}. \end{aligned}$$

Hence $\liminf_{n \rightarrow \infty} \langle m, f_n \rangle \geq r + \langle m_a, f \rangle$.

Now since the function $\pi(\xi(x) + t)$ is a convex function of $t \in \mathbb{R}$, we have for each $x \in X$

$$\pi(\xi(x) + f_n(x)) \leq \tau_n(x)\pi(\xi(x) + g(x)) + (1 - \tau_n(x))\pi(\xi(x) + f(x)).$$

Therefore

$$\begin{aligned} C(f_n) &\leq \int_{O_n} \tau_n(x) \{ \pi(\xi(x) + g(x)) - \pi(\xi(x) + f(x)) \} \mu(dx) + C(f) \\ &\leq \int_{O_n} |\pi(\xi(x) + g(x))| \mu(dx) + \int_{O_n} |\pi(\xi(x) + f(x))| \mu(dx) + C(f). \end{aligned}$$

Thus by condition (ii) of Hypothesis 2 and Lebesgue's Dominated Convergence Theorem

$$\limsup_{n \rightarrow \infty} C(f_n) \leq C(f).$$

It follows that

$$I(m) \geq \liminf_{n \rightarrow \infty} (\langle m, f_n \rangle - C(f_n)) \geq r + \langle m_a, f \rangle - C(f)$$

and so $I(m) \geq r + I(m_a)$. Since r is an arbitrary number less than $U(m_s)$, we get

$$I(m) \geq U(m_s) + I(m_a).$$

This completes the proof. □

For use in the next section, we note the following simple corollary of Lemma 2.7.

Corollary 2.8 *If $m \in E$ satisfies $I(m) < \infty$ and if $\gamma = \infty$, then m_s , the singular part of m relative to μ , equals 0.*

Proof: By Lemma 2.7, we must have $U(m_s) < \infty$. If $m_s \neq 0$, then since $\gamma = \infty$, we would have $U(m_s) = \infty$. We conclude that $m_s = 0$. □

In the next theorem we give an explicit form of the rate function I in the large deviation principle. Let $\pi^* : \mathbb{R} \mapsto (-\infty, \infty]$ be the Legendre-Fenchel transform of π ; that is

$$\pi^*(t) \doteq \sup_{s < \gamma} (ts - \pi(s)).$$

For $t \in \mathbb{R}$ and $r < \gamma$ let

$$J(t, r) \doteq \pi^*(t) - rt + \pi(r).$$

Note that $J(t, r) \geq 0$ and that $t \mapsto J(t, r)$ is lower semicontinuous.

Theorem 3. For each $m \in E$, let m_a and m_s be respectively the absolutely continuous part and the singular part of m in the Lebesgue decomposition of m relative to μ . Then

$$I(m) = U(m_s) + \int_X J(\rho(x), \xi(x)) \mu(dx)$$

where $\rho(x) \doteq \frac{dm_a}{d\mu}(x)$.

Proof: By Lemma 2.7, $I(m) = U(m_s) + I(m_a)$. Hence we must prove that

$$I(m_a) = \int_X J(\rho(x), \xi(x)) \mu(dx)$$

Define $J(t, r; s) = ts - \pi(s) - rt + \pi(r)$ for $t \in \mathbb{R}$, $r < \gamma$ and $s < \gamma$, so that

$$J(t, r) = \sup_{s < \gamma} J(t, r; s).$$

If $f \in \mathcal{D}$ then

$$\begin{aligned} \langle m_a, f \rangle - C(f) &= \langle m_a, f \rangle - \int_X \{\pi(\xi(x) + f(x)) - \pi(\xi(x))\} \mu(dx) \\ &= \int_X f(x) \rho(x) \mu(dx) - \int_X \{\pi(\xi(x) + f(x)) - \pi(\xi(x))\} \mu(dx) \\ &= \int_X J(\rho(x), \xi(x), f(x) + \xi(x)) \mu(dx) \\ &\leq \int_X J(\rho(x), \xi(x)) \mu(dx). \end{aligned}$$

Therefore $I(m_a) \leq \int_X J(\rho(x), \xi(x)) \mu(dx)$.

In order to complete the proof we must show the opposite inequality. If $I(m_a) = \infty$, there is nothing more to prove. Suppose $I(m_a) < \infty$. By Theorem 1(a) there is a sequence $\{f_n\}$ in \mathcal{D} such that if $\rho_n = \rho^{f_n}$, then ρ_n converges to ρ in the L^1 -norm with respect to μ and if $m^{f_n}(dx) = \rho_n(x) \mu(dx)$, then $\lim_{n \rightarrow \infty} I(m^{f_n}) = I(m_a)$. The sequence $\{\rho_n\}$ has a subsequence $\{\rho_{n_k}\}$ which converges pointwise ρ , μ -a.e. Now since $t \mapsto J(t, r)$ is lower semi-continuous we have

$$\begin{aligned} \int_X J(\rho(x), \xi(x)) \mu(dx) &\leq \int_X \liminf_{k \rightarrow \infty} J(\rho_{n_k}(x), \xi(x)) \mu(dx) \\ &\leq \liminf_{k \rightarrow \infty} \int_X J(\rho_{n_k}(x), \xi(x)) \mu(dx) \end{aligned}$$

by Fatou's Lemma. But by Lemma 2.4

$$\begin{aligned}
I(m^{f_n}) &= I(m^{f_n}, f_n) \\
&= \int_X f_n(x) m^{f_n}(dx) - C(f_n) \\
&= \int_X \{f_n(x) \rho_n(x) + \pi(\xi(x)) - \pi(\xi(x) + f_n(x))\} \mu(dx) \\
&= \int_X \{[(\xi(x) + f_n(x)) \rho_n(x) - \pi(\xi(x) + f_n(x))] - \xi(x) \rho_n(x) + \pi(\xi(x))\} \mu(dx).
\end{aligned}$$

Since for $s < \gamma$, $\pi^*(\pi'(s)) = s\pi'(s) - \pi(s)$ and $\rho_n(x) = \pi'(\xi(x) + f_n(x))$,

$$\pi^*(\rho_n(x)) = (\xi(x) + f_n(x)) \rho_n(x) - \pi(\xi(x) + f_n(x)).$$

Therefore

$$\begin{aligned}
I(m^{f_n}) &= \int_X [(\pi^*(\rho_n(x)) - \xi(x) \rho_n(x) + \pi(\xi(x)))] \mu(dx) \\
&= \int_X J(\rho_n(x), \xi(x)) \mu(dx).
\end{aligned}$$

Thus we have

$$\int_X J(\rho(x), \xi(x)) \mu(dx) \leq \liminf_{k \rightarrow \infty} I(m^{f_{n_k}}) = I(m_a).$$

This completes the proof. □

In the next section we prove the Approximation Theorem, Theorem 1.

§3. Proof of the Approximation Theorem

In this section we shall prove Theorem 1. We first prove Theorem 1 (a) and then show that Theorem 1 (b) is a corollary of (a). In part (a) we want to approximate the density ρ of a measure m that is absolutely continuous with respect to μ by the density ρ^f of a measure m^f with $f \in \mathcal{D}$. We recall that

$$\rho^f(x) \doteq \pi'(\xi(x) + f(x)),$$

so that the range of ρ^f must be a subset of the range of π' . In the first lemma, Lemma 3.1, we prove that if $m \in E$ is such that $I(m)$ is finite and m is absolutely continuous with respect to μ with density ρ , then $\rho(x)$ must be within the range of π' almost everywhere with respect to μ . The idea of the proof is that if $t \in \mathbb{R}$ is outside the range of π' , then the Legendre-Fenchel transform $\pi^*(t)$ is infinite.

Let the range of π' be (ρ_1, ρ_2) ; i.e. $\rho_1 \doteq \lim_{s \rightarrow -\infty} \pi'(s)$ and $\rho_2 \doteq \lim_{s \uparrow \gamma} \pi'(s)$. These limits exist since π' is monotonic.

Lemma 3.1 *Let $m \in E$ and $\rho(x) \doteq \frac{dm_a}{d\mu}(x)$, where m_a is the absolutely continuous part of m in the decomposition relative to μ . If $\mu(\{x \in X : \rho(x) > \rho_2\} \cup \{x \in X : \rho(x) < \rho_1\}) \neq 0$, then $I(m) = \infty$.*

Proof: For $s, s_1 \in (-\infty, \gamma)$ we have

$$\pi(s) - \pi(s_1) \leq (s - s_1)\pi'(s),$$

and therefore for $\rho \in \mathbb{R}$,

$$\rho s - \pi(s) \geq (\rho - \pi'(s))s + s_1\pi'(s) - \pi(s_1). \quad (3.1)$$

Let $s_1 < 0 \wedge \gamma$. Then for $\rho < \rho_1$ and $s < s_1$

$$\rho s - \pi(s) \geq (\rho - \rho_1)s + s_1\pi'(s_1) - \pi(s_1). \quad (3.2)$$

Let $C \doteq \{x \in X : \rho(x) < \rho_1\}$ and suppose $\mu(C) > 0$. Let $C_0 \doteq \{x \in X : \rho(x) < \rho_1 - 1\}$ and for $n = 1, 2, \dots$ let

$$C_n \doteq \{x \in X : \rho_1 - \frac{1}{n+1} > \rho(x) \geq \rho_1 - \frac{1}{n}\}.$$

Then $C = \bigcup_{n \geq 0} C_n$ and therefore for some n , $\mu(C_n) > 0$. Thus there is an $\epsilon > 0$ such that $\mu\{x \in X : \rho(x) < \rho_1 - \epsilon\} > 0$. Since μ is a regular measure, we can also then find a compact set $K \subset \{x : \rho(x) < \rho_1 - \epsilon\}$ such that $\mu(K) > 0$. Let

$c_1 \doteq \inf_{x \in K} \xi(x)$ and $c_2 \doteq \sup_{x \in K} \xi(x)$ and choose $s < \min(s_1, c_1, \frac{s_1}{\epsilon}, \pi'(s_1))$. Again since μ is regular, we can then find an open set O such that $K \subset O$ and $m(O \setminus K) < \frac{1}{c_2 - s}$. By Urysohn's Lemma we can find $\tau \in \mathcal{C}^b(X)$ such that $0 \leq \tau(x) \leq 1$ for all $x \in X$, $\tau(x) = 1$ for $x \in K$ and $\tau(x) = 0$ for $x \in O^c$. Let

$$g(x) \doteq \begin{cases} s - c_2 & \xi(x) > c_2, \\ s - \xi(x) & c_1 \leq \xi(x) \leq c_2, \\ s - c_1 & \xi(x) < c_1, \end{cases}$$

and let $f(x) \doteq \tau(x)g(x)$; then $f \in \mathcal{D}$. Since g satisfies $s - c_2 \leq g(x) \leq s - c_1 < 0$ for all $x \in X$, f satisfies

$$s - c_2 \leq f(x) \leq 0$$

for all $x \in X$. By Lemma 2.7, $I(m) = U(m_s) + I(m_a) \geq I(m_a)$. (Note that the proof of Lemma 2.7 does not use Theorem 1.) Now

$$\begin{aligned} I(m_a) &\geq I(m_a, f) = \int_K \{\rho(x)(f(x) + \xi(x)) - \pi(f(x) + \xi(x))\} \mu(dx) \\ &\quad - \int_K \{\rho(x)\xi(x) - \pi(\xi(x))\} \mu(dx) \\ &\quad + \int_{O \setminus K} \rho(x)f(x) + \int_{O \setminus K} \{\pi(\xi(x)) - \pi(f(x) + \xi(x))\} \mu(dx). \end{aligned} \quad (3.3)$$

Since $f(x) \leq 0$,

$$\begin{aligned} I(m_a) &\geq \int_K \{\rho(x)(f(x) + \xi(x)) - \pi(f(x) + \xi(x))\} \mu(dx) \\ &\quad - \int_K \{\rho(x)\xi(x) - \pi(\xi(x))\} \mu(dx) + \int_{O \setminus K} \rho(x)f(x). \end{aligned}$$

Thus by (3.2), since $f(x) + \xi(x) = s$ for $x \in K$

$$\begin{aligned} I(m_a) &\geq (\epsilon|s| - |s_1|\pi'(s_1) - \pi(s_1))\mu(K) - c_2\|m\| + (s - c_2)m(O \setminus K) \\ &\quad - \int_X |\pi(\xi(x))| \mu(dx) \\ &= (\epsilon|s| - |s_1|\pi'(s_1) - \pi(s_1))\mu(K) - c_2\|m\| - 1 - \int_X |\pi(\xi(x))| \mu(dx). \end{aligned}$$

Letting $s \rightarrow -\infty$ we get $I(m_a) = \infty$ and therefore $I(m) = \infty$.

Suppose $\mu\{x \in X : \rho(x) > \rho_2\} > 0$. Clearly this is not possible if $\rho_2 = \infty$. Therefore we can assume $\rho_2 < \infty$. One of the assumptions in Section 1 was that if $\gamma < \infty$, then $\rho_2 = \infty$. Therefore here we can take $\gamma = \infty$. Then from (3.1) with $s_1 = 0$, we get for $s \geq 0$

$$\rho s - \pi(s) \geq (\rho - \rho_2)s - \pi(0). \quad (3.4)$$

By an argument similar to the above we can choose a compact set $K \subset \{x : \rho(x) > \rho_2 + \epsilon\}$ such that $\mu(K) > 0$. With the same definitions of c_1 and c_2 we let $s > (0 \vee c_2)$. Since the measure $|\pi(\xi(x) + s - c_1)|\mu(dx)$ is regular, we can find an open set O such that $K \subset O$ and $\int_{O \setminus K} |\pi(\xi(x) + s - c_1)|\mu(dx) < 1$. Let g and f be defined as above. In this case $0 \leq s - c_2 \leq g(x) \leq s - c_1$ and thus

$$0 \leq f(x) \leq s - c_1 \quad (3.5)$$

for all $x \in X$. Using (3.3) and the fact that $f(x) \geq 0$, we get

$$\begin{aligned} I(m_a) &\geq \int_K \{\rho(x)(f(x) + \xi(x)) - \pi(f(x) + \xi(x))\} \mu(dx) \\ &= - \int_K \{\rho(x)\xi(x) - \pi(\xi(x))\} \mu(dx) + \int_{O \setminus K} \{\pi(\xi(x)) - \pi(f(x) + \xi(x))\} \mu(dx). \end{aligned}$$

Thus by (3.4) we get

$$\begin{aligned} I(m_a) &\geq (\epsilon s - \pi(0))\mu(K) - c_2 \|m\| - \int_{O \setminus K} \pi(\xi(x) + s - c_1) \mu(dx) + \int_O \pi(\xi(x)) \mu(dx) \\ &\geq (\epsilon s - \pi(0))\mu(K) - c_2 \|m\| - 1 - \int_X |\pi(\xi(x))| \mu(dx). \end{aligned}$$

Letting $s \rightarrow \infty$ we obtain $I(m_a) = \infty$ and hence $I(m) = \infty$. The proof of the lemma is complete. □

Lemma 3.2 *If $m \in E$ and $I(m) < \infty$, then $\int_X |\xi(x)| m(dx) < \infty$.*

Proof: We prove that $\int_X |\xi(x)| m(dx) = \infty$ implies $I(m) = \infty$. Suppose first that $\gamma < \infty$. If $\int |\xi(x)| m(dx) = \infty$, then since

$$\gamma - \xi(x) = |\gamma - \xi(x)| \geq |\xi(x)| - |\gamma|,$$

we have $\int_X (\gamma - \xi(x)) m(dx) = \infty$. If $\int_X (\gamma - \xi(x)) m_s(dx) = \infty$, then $I(m) = \infty$ by Lemma 2.7 since $I(m_a) \geq 0$. Suppose $\int_X (\gamma - \xi(x)) m_a(dx) = \infty$. Then since $(\gamma - \xi(x)) m_a(dx)$ is a Radon measure, given $r \in \mathbb{R}$ there exists K , a compact subset of X , such that

$$\int_K (\gamma - \xi(x)) m_a(dx) > r.$$

Define $f : X \mapsto \mathbb{R}$ by

$$f(x) = \left\{ \frac{1}{2}(\gamma - \xi(x)) \right\} \wedge \left\{ \frac{1}{2}(\gamma - \inf_{y \in K} \xi(y)) \right\}.$$

Then

$$0 \leq f(x) \leq \frac{1}{2}(\gamma - \inf_{y \in K} \xi(y))$$

for $x \in X$ and

$$\xi(x) + f(x) < \frac{1}{2}(\gamma + \xi(x)) < \frac{1}{2}(\gamma + \xi_0)$$

for all $x \in X$; hence $f \in \mathcal{D}$. Also $f(x) = \frac{1}{2}(\gamma - \xi(x))$ for $x \in K$. Thus

$$\begin{aligned} I(m_a) &\geq I(m_a, f) \geq \frac{1}{2} \int_K (\gamma - \xi(x)) m_a(dx) - \int_X \{\pi(\xi(x) + f(x)) - \pi(\xi(x))\} \mu(dx) \\ &\geq r - \int_X \{|\pi(\frac{1}{2}(\gamma + \xi(x)))| + |\pi(\xi(x))|\} \mu(dx). \end{aligned}$$

Since r is arbitrary, $I(m_a) = \infty$. But $U(m_s) \geq 0$, and so by Lemma 2.7 $I(m) = \infty$.

Suppose that $\gamma = \infty$. By Corollary 2.8, if $m_s \neq 0$, then $I(m) = \infty$. (Note that the proof of Corollary 2.8 does not use Theorem 1.) Therefore we can assume that $m_s = 0$. If $\int_X |\xi(x)| m(dx) = \infty$, then at least one of the quantities $\int_X \xi_+(x) m(dx)$ and $\int_X \xi_-(x) m(dx)$ is $+\infty$. Suppose $\int_X \xi_+(x) m(dx) = \infty$. Given $r \in \mathbb{R}$ there exists $K \subset X$, a compact subset of X , such that $\int_K \xi_+(x) m(dx) > r$. Define

$$f(x) \doteq \{\xi_+(x)\} \wedge \left\{ \sup_{y \in K} \xi_+(y) \right\}.$$

Then $f \in \mathcal{D}$ and

$$\begin{aligned} I(m) &\geq I(m, f) \geq r - \int_{\{x: \xi(x) \geq 0\}} \{\pi(2\xi(x)) - \pi(\xi(x))\} \mu(dx) \\ &\geq r - \int_X |\pi(2\xi(x))| - \int_X |\pi(\xi(x))| \mu(dx), \end{aligned}$$

and thus since r is arbitrary, $I(m) = \infty$. Similarly, if $\int_X \xi_-(x) m(dx) = \infty$, given $r \in \mathbb{R}$ there exists $K \subset X$, a compact subset of X , such that $\int_K \xi_-(x) m(dx) > r$. The same argument works with the function

$$f(x) \doteq \frac{1}{2} \{\xi_-(x)\} \wedge \left\{ \sup_{y \in K} \xi_-(y) \right\}.$$

In this case

$$\begin{aligned} I(m) &\geq I(m, f) \geq r - \int_{\{x: \xi(x) \leq 0\}} \left\{ \pi\left(\frac{1}{2}\xi(x)\right) - \pi(\xi(x)) \right\} \mu(dx) \\ &\geq r - \int_X \left| \pi\left(\frac{1}{2}\xi(x)\right) \right| \mu(dx) - \int_X |\pi(\xi(x))| \mu(dx), \end{aligned}$$

and again, since r is arbitrary, $I(m) = \infty$.

□

We shall now proceed with the proof of Theorem 1 (a).

Proof of Theorem 1 (a):

Let $m \in E$ be absolutely continuous with respect to μ and let $\rho(x) = \frac{dm}{d\mu}(x)$.

We first treat the case when $\rho_1 = \rho_2 = \rho_0$, say. In this case by Lemma 3.1 $I(m) < \infty$ implies that $\rho(x) = \rho_0$ μ -a.e. On the other hand we must then have $\pi'(\xi(x)) = \rho_0$ for all $x \in X$. Therefore $\rho(x) = \rho^f(x)$ μ -a.e. with $f(x) = 0$ and Theorem 1 (a) is immediate.

We now suppose that $0 \leq \rho_1 < \rho_2$. Then σ does not consist of a single atom, for otherwise $\pi'(s)$ would be constant. If σ does not consist of a single atom, then $\pi''(s) > 0$ and π' is strictly increasing. Therefore π' is invertible on (ρ_1, ρ_2) .

Let $f(x) \doteq (\pi')^{-1}(\rho(x)) - \xi(x)$. If $f \in \mathcal{D}$, the conclusion in Theorem 1 (a) is true since then $\rho = \rho^f$. However in general $f \notin \mathcal{D}$ and we have to make some approximations. We first approximate ρ by a continuous function $\tilde{\rho}_n$ since we want $\tilde{f}_n(x) \doteq (\pi')^{-1}(\tilde{\rho}_n(x)) - \xi(x)$ to be continuous. Also we have to trim $\tilde{\rho}_n$ so that $\tilde{\rho}_n(x)$ is in the range of π' .

For $n \in \mathbb{N}$, $n > (\rho_2 - \rho_1)^{-1}$ let

$$A_n \doteq \{n \vee |(\pi')^{-1}(n \wedge (\rho_2 - n^{-1}))|\}.$$

Since $I(m) < \infty$, we have by Lemma 3.2

$$\int_X |\xi(x)| m(dx) < \infty.$$

Define the measure

$$\tilde{\mu}(dx) \doteq (1 + |\xi(x)|) \mu(dx).$$

Since $\int_X \rho(x) \tilde{\mu}(dx) < \infty$, for each $n \in \mathbb{N}$ we can find $\tilde{\rho}_n \in C(X)$ with compact support such that

$$\int_X |\tilde{\rho}_n(x) - \rho(x)| \tilde{\mu}(dx) < A_n^{-2}$$

(see for example Rudin [8], Theorem 3.14). Then it follows that

$$\int_X |\tilde{\rho}_n(x) - \rho(x)| \mu(dx) < A_n^{-2}$$

and

$$\int_X |\tilde{\rho}_n(x) - \rho(x)| |\xi(x)| \mu(dx) < A_n^{-2}.$$

Now we trim $\tilde{\rho}_n$ in a suitable way. For each $n \in \mathbb{N}$, $n > (\rho_2 - \rho_1)^{-1}$ and each $x \in X$ let $\beta_n(x) \doteq \pi'(\xi(x) - n)$ and $\gamma_n(x) \doteq (\pi'(\xi(x) - n) \vee (n \wedge [\rho_2 - n^{-1}]))$. Then $0 \leq \rho_1 < \beta_n(x) \leq \gamma_n(x) < \rho_2$ for all $x \in X$. Let

$$\begin{aligned} E_n^< &\doteq \{x \in X : \tilde{\rho}_n(x) < \beta_n(x)\}, \\ E_n^0 &\doteq \{x \in X : \beta_n(x) \leq \tilde{\rho}_n(x) \leq \gamma_n(x)\}, \\ E_n^> &\doteq \{x \in X : \tilde{\rho}_n(x) > \gamma_n(x)\}. \end{aligned}$$

Define $f_n \in \mathcal{D}$ by

$$f_n(x) \doteq \begin{cases} (\pi')^{-1}(\beta_n(x)) - \xi(x) = -n, & x \in E_n^<, \\ (\pi')^{-1}(\tilde{\rho}_n(x)) - \xi(x), & x \in E_n^0, \\ (\pi')^{-1}(\gamma_n(x)) - \xi(x), & x \in E_n^>. \end{cases}$$

Note that since $\tilde{\rho}_n$ has compact support and $f_n(x) = -n$ for $x \notin \text{supp } \tilde{\rho}_n$, f_n is bounded. Also it is easy to check that

$$|f_n(x)| \leq |\xi(x)| + A_n$$

for all $x \in X$. Let $\rho_n \doteq \rho^{f_n}$; that is, let

$$\rho_n(x) \doteq \begin{cases} \beta_n(x) & x \in E_n^<, \\ \tilde{\rho}_n(x) & x \in E_n^0, \\ \gamma_n(x) & x \in E_n^>. \end{cases}$$

Let $m_n \doteq m^{f_n}$; that is $m_n(dx) \doteq \rho_n(x) \mu(dx)$. We want to show that the trimmed density ρ_n is still a good approximation to ρ ; that is, we want to prove that $\int_X |\rho(x) - \rho_n(x)| \mu(x)$ tends to zero as n tends to infinity. Now

$$\int_X |\rho(x) - \rho_n(x)| \mu(x) = \int_{E_n^0} |\rho(x) - \tilde{\rho}_n(x)| \mu(dx) + \int_{E_n^< \cup E_n^>} |\rho(x) - \rho_n(x)| \mu(dx).$$

Also

$$\begin{aligned}
& \int_{E_n^<} |\rho(x) - \rho_n(x)| \mu(dx) \\
&= \int_{E_n^<} |\rho(x) - \beta_n(x)| \mu(dx) \\
&= \int_{\{x: \tilde{\rho}_n(x) \vee \rho(x) < \beta_n(x)\}} |\rho(x) - \beta_n(x)| \mu(dx) + \int_{\{x: \tilde{\rho}_n(x) < \beta_n(x) < \rho(x)\}} |\rho(x) - \beta_n(x)| \mu(dx) \\
&= \int_{\{x: \tilde{\rho}_n(x) \vee \rho(x) < \beta_n(x)\}} (\beta_n(x) - \rho(x)) \mu(dx) + \int_{\{x: \tilde{\rho}_n(x) < \beta_n(x) < \rho(x)\}} (\rho(x) - \beta_n(x)) \mu(dx) \\
&\leq \int_{\{x: \rho(x) < \beta_n(x)\}} (\beta_n(x) - \rho(x)) \mu(dx) + \int_{\{x: \tilde{\rho}_n(x) < \beta_n(x) < \rho(x)\}} |\rho(x) - \tilde{\rho}_n(x)| \mu(dx) \\
&\leq \int_{\{x: \rho(x) < \beta_n(x)\}} (\beta_n(x) - \rho(x)) \mu(dx) + \int_{E^<} |\rho(x) - \tilde{\rho}_n(x)| \mu(dx).
\end{aligned}$$

Similarly

$$\begin{aligned}
& \int_{E_n^>} |\rho(x) - \rho_n(x)| \mu(dx) \leq \int_{E^>} |\rho(x) - \tilde{\rho}_n(x)| \mu(dx) \\
&+ \int_{\{x: \rho(x) > n \wedge (\rho_2 - n^{-1})\}} (\rho(x) - n \wedge (\rho_2 - n^{-1})) \mu(dx).
\end{aligned}$$

Thus

$$\begin{aligned}
& \int_X |\rho(x) - \rho_n(x)| \mu(x) \leq \int_X |\rho(x) - \tilde{\rho}_n(x)| \mu(dx) \\
&+ \int_{\{x: \rho(x) < \pi'(\xi(x) - n)\}} (\pi'(\xi(x) - n) - \rho(x)) \mu(dx) \\
&+ \int_{\{x: \rho(x) > n \wedge (\rho_2 - n^{-1})\}} (\rho(x) - n \wedge (\rho_2 - n^{-1})) \mu(dx).
\end{aligned}$$

We consider the three integrals in the last display separately. For the first integral, we have

$$\int_X |\rho(x) - \tilde{\rho}_n(x)| \mu(dx) < A_n^{-2} \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Since by Lemma 3.1 $\mu\{x : \rho(x) < \rho_1\} = 0$, we have for the second integral

$$\begin{aligned} \int_{\{x: \rho(x) < \pi'(\xi(x)-n)\}} \{\pi'(\xi(x)-n) - \rho(x)\} \mu(dx) \leq \\ \int_{\{x: \rho(x) = \rho_1\}} \{\pi'(\xi(x) - n) - \rho_1\} \mu(dx) + \int_{\{x: \rho_1 < \rho(x) < \pi'(\xi(x)-n)\}} \pi'(\xi(x)) \mu(dx). \end{aligned}$$

Both terms tend to zero as $n \rightarrow \infty$ by Lebesgue's Dominated Convergence Theorem.

For the third integral we have to consider the cases $\rho_2 = \infty$ and $\rho_2 < \infty$ separately. If $\rho_2 = \infty$, then

$$\begin{aligned} \int_{\{x: \rho(x) > n \wedge (\rho_2 - n^{-1})\}} (\rho(x) - n \wedge (\rho_2 - n^{-1})) \mu(dx) &= \int_{\{x: \rho(x) > n\}} (\rho(x) - n) \mu(dx) \\ &\leq \int_{\{x: \rho(x) > n\}} \rho(x) \mu(dx), \end{aligned}$$

which tends to zero by Lebesgue's Dominated Convergence Theorem. If $\rho_2 < \infty$, since by Lemma 3.1 $\mu\{x : \rho(x) > \rho_2\} = 0$, we have for all sufficiently large n

$$\int_{\{x: \rho(x) > n \wedge (\rho_2 - n^{-1})\}} (\rho(x) - n \wedge (\rho_2 - n^{-1})) \mu(dx) \leq \frac{1}{n\rho_2} \int_X \rho(x) \mu(dx) + \int_{\{x: \rho_2 > \rho(x) > \rho_2 - \frac{1}{n}\}} \rho(x) \mu(dx).$$

Again these integrals tend to zero, the second by Lebesgue's Dominated Convergence Theorem. Therefore

$$\lim_{n \rightarrow \infty} \int_X |\rho(x) - \rho_n(x)| \mu(x) = 0.$$

It is easy to deduce now that $m_n \rightarrow m$ in the narrow topology. For $g \in C^b(X)$,

$$|\langle m, g \rangle - \langle m_n, g \rangle| = \left| \int_X g(x) (\rho_n(x) - \rho(x)) \mu(dx) \right| \leq \|g\|_\infty \int_X |\rho_n(x) - \rho(x)| \mu(dx).$$

Therefore $m_n \rightarrow m$ in the narrow topology.

To complete the proof of Theorem 1 (a) we have to prove that $\lim_{n \rightarrow \infty} I(m_n) = I(m)$. Since $m_n = m^{f_n}$, we have by Lemma 2.4

$$\begin{aligned} I(m_n) &= \langle m_n, f_n \rangle - C(f_n) \\ &= \langle m_n, f_n \rangle - \langle m, f_n \rangle + I(m, f_n) \\ &\leq \langle m_n, f_n \rangle - \langle m, f_n \rangle + I(m). \end{aligned}$$

Thus if $\limsup_{n \rightarrow \infty} (\langle m_n, f_n \rangle - \langle m, f_n \rangle) \leq 0$, then $\limsup_{n \rightarrow \infty} I(m_n) \leq I(m)$. But since $m_n \rightarrow m$ in the narrow topology and I is lower semi-continuous, we also have

$$\liminf_{n \rightarrow \infty} I(m_n) \geq I(m).$$

Therefore $\lim_{n \rightarrow \infty} I(m_n) = I(m)$. To finish the proof it is sufficient to show that

$$\limsup_{n \rightarrow \infty} (\langle m_n, f_n \rangle - \langle m, f_n \rangle) \leq 0.$$

We have

$$\begin{aligned} \langle m_n, f_n \rangle - \langle m, f_n \rangle &= \int_X f_n(x)(\rho_n(x) - \rho(x))\mu(dx) \\ &= \int_X f_n(x)(\rho_n(x) - \tilde{\rho}_n(x))\mu(dx) \\ &\quad + \int_X f_n(x)(\tilde{\rho}_n(x) - \rho(x))\mu(dx) \\ &\leq \int_{E_n^< \cup E_N^>} f_n(x)(\rho_n(x) - \tilde{\rho}_n(x))\mu(dx) \\ &\quad + \int_X |f_n(x)| |\tilde{\rho}_n(x) - \rho(x)| \mu(dx). \end{aligned}$$

Also

$$\begin{aligned} \int_X |f_n(x)| |\tilde{\rho}_n(x) - \rho(x)| \mu(dx) &\leq \int_X |\xi(x)| |\tilde{\rho}_n(x) - \rho(x)| \mu(dx) \\ &+ A_n \int_X |\tilde{\rho}_n(x) - \rho(x)| \mu(dx) \leq A_n^{-2} + A_n^{-1} \rightarrow 0 \text{ as } n \rightarrow \infty. \end{aligned}$$

For $x \in E_n^<$, $f_n(x) = -n$ and $\rho_n(x) > \tilde{\rho}_n(x)$ and therefore

$$\int_{E_n^<} f_n(x)(\rho_n(x) - \tilde{\rho}_n(x))\mu(dx) \leq 0.$$

Let $r_n \doteq (\pi')^{-1}(n \wedge (\rho_2 - n^{-1}))$. Clearly there is $r \in \mathbb{R}$ such that $r_n > r$ for all n .

Let $x \in E_N^>$ and suppose $\pi'(\xi(x) - n) \geq n \wedge (\rho_2 - n^{-1})$. Then $\xi(x) - n \geq r_n$ and

$$f_n(x) = -n \geq r_n - \xi(x) > r - \xi(x).$$

If $x \in E_n^>$ and $\pi'(\xi(x) - n) < n \wedge (\rho_2 - n^{-1})$, then $f_n(x) = r_n - \xi(x) > r - \xi(x)$.
Now for $x \in E_n^>$, $\rho_n(x) < \tilde{\rho}_n(x)$ and therefore

$$\begin{aligned}
& \int_{E_n^>} f_n(x)(\rho_n(x) - \tilde{\rho}_n(x))\mu(dx) \\
& < \int_{E_n^>} (r - \xi(x))(\rho_n(x) - \tilde{\rho}_n(x))\mu(dx) \\
& \leq \int_{E_n^>} (|r| + |\xi(x)|)|\rho_n(x) - \tilde{\rho}_n(x)|\mu(dx) \\
& \leq \int_{E_n^>} (|r| + |\xi(x)|)|\rho(x) - \rho_n(x)|\mu(dx) + (|r| + 1)A_n^{-2} \\
& \leq \int_{E_n^> \cap \{x: \rho(x) > \gamma_n(x)\}} (|r| + |\xi(x)|)(\rho(x) - \gamma_n(x))\mu(dx) \\
& \quad + \int_{E_n^> \cap \{x: \rho(x) \leq \gamma_n(x)\}} (|r| + |\xi(x)|)(\tilde{\rho}_n(x) - \rho(x))\mu(dx) + (|r| + 1)A_n^{-2} \\
& \leq 2(|r| + 1)A_n^{-2} + \int_{E_n^> \cap \{x: \rho(x) > \gamma_n(x)\}} (|r| + |\xi(x)|)(\rho(x) - \gamma_n(x)) \\
& \leq 2(|r| + 1)A_n^{-2} + \int_{\{x: \rho(x) > n \wedge (\rho_2 - n^{-1})\}} (|r| + |\xi(x)|)(\rho(x) - n \wedge (\rho_2 - n^{-1}))\mu(dx).
\end{aligned}$$

If $\rho_2 = \infty$, then

$$\begin{aligned}
& \int_{\{x: \rho(x) > n \wedge (\rho_2 - n^{-1})\}} (|r| + |\xi(x)|)(\rho(x) - n \wedge (\rho_2 - \frac{1}{n}))\mu(dx) \\
& = \int_{\{x: \rho(x) > n\}} (|r| + |\xi(x)|)(\rho(x) - n)\mu(dx) \leq \int_{\{x: \rho(x) > n\}} (|r| + |\xi(x)|)\rho(x)\mu(dx)
\end{aligned}$$

which tends to zero by Lebesgue's Dominated Convergence Theorem. If $\rho_2 < \infty$, since $\rho_2 > 0$,

$$\begin{aligned}
& \int_{\{x: \rho(x) > n \wedge (\rho_2 - n^{-1})\}} (|r| + |\xi(x)|)(\rho(x) - n \wedge (\rho_2 - n^{-1}))\mu(dx) \\
& \leq \frac{1}{n} \int_{\{x: \rho(x) = \rho_2\}} (|r| + |\xi(x)|)\mu(dx) + \int_{\{x: \rho_2 > \rho(x) > \rho_2 - \frac{1}{n}\}} (|r| + |\xi(x)|)\rho(x)\mu(dx) \\
& \leq \frac{1}{n\rho_2} \int_X (|r| + |\xi(x)|)\mu(dx) + \int_{\{x: \rho_2 > \rho(x) > \rho_2 - \frac{1}{n}\}} (|r| + |\xi(x)|)\rho(x)\mu(dx).
\end{aligned}$$

Both terms tends to zero as $n \rightarrow \infty$ and thus $\limsup_{n \rightarrow \infty} (\langle m_n, f_n \rangle - \langle m, f_n \rangle) \leq 0$. This completes the proof of Theorem 1 (a). □

We now turn to the proof of Theorem 1 (b). The proof proceeds by a sequence of reductions. First we show (Lemma 3.3) that if $m \in E$ and $I(m) < \infty$, then m and $I(m)$ can be approximated by $m^{(n)} \in E$ and $I(m^{(n)})$ respectively, where the singular part in the decomposition of $m^{(n)}$ relative to μ has compact support. We then show (Lemma 3.5) that analogous approximations can be made using a measure m' that is absolutely continuous with respect to μ . Then Theorem 1, part (b) follows from part (a). Lemma 3.4 is used in Lemma 3.5 in approximating $I(m)$ by $I(m')$.

Lemma 3.3 *If $m \in E$ and $I(m) < \infty$, then there is a sequence $\{m^{(n)}\}$ in E such that $m^{(n)}$ converges to m , $\lim_{n \rightarrow \infty} I(m^{(n)}) = I(m)$, and the singular part in the decomposition of each $m^{(n)}$ relative to μ , $m_s^{(n)}$, has compact support for each n .*

Proof: If $\gamma = \infty$, then by Corollary 2.8 $m_s = 0$. In this case we set $m^{(n)} = m$ for all n and we are done. Now suppose $\gamma < \infty$. Let $\{K_n\}$ be a sequence of compact subsets of X such that $K_n \subset K_{n+1}$ and $m(X \setminus K_n) < \frac{1}{n}$. We can find such a sequence $\{K_n\}$ because m is a Radon measure. Let $\tilde{m}^{(n)}(dx) \doteq 1_{K_n}(x)m_s(dx)$. If $f \in C^b(X)$, then

$$|\langle \tilde{m}^{(n)}, f \rangle - \langle m_s, f \rangle| \leq \frac{1}{n} \|f\|_\infty.$$

Therefore $\tilde{m}^{(n)} \rightarrow m_s$ in the narrow topology and consequently $m^{(n)} \doteq \tilde{m}^{(n)} + m_a$ converges to m in the narrow topology. Now

$$U(m_s^{(n)}) = U(\tilde{m}^{(n)}) = \int_X (\gamma - \xi(x)) \tilde{m}^{(n)}(dx) = \int_X (\gamma - \xi(x)) 1_{K_n}(x) m_s(dx).$$

Thus by Lebesgue's Monotone Convergence Theorem

$$\lim_{n \rightarrow \infty} (U(m_s^{(n)})) = \int_X (\gamma - \xi(x)) m_s(dx) = U(m_s).$$

Therefore by Lemma 2.7

$$\lim_{n \rightarrow \infty} I(m^{(n)}) = \lim_{n \rightarrow \infty} U(m_s^{(n)}) + I(m_a) = U(m_s) + I(m_a) = I(m).$$

This completes the proof. □

Lemma 3.4 *Let m be a measure in E having compact support. If N is a neighbourhood of m , $f \in \mathcal{C}(X)$, and $\epsilon > 0$, then there exists $m' \in N$ such that m' is absolutely continuous with respect to μ and satisfies*

$$\left| \int_X f(x)m(dx) - \int_X f(x)m'(dx) \right| < \epsilon.$$

Proof: Let $K \doteq \text{supp } m$. There are $f_1 \dots f_n \in \mathcal{C}^b(X)$ and $\delta > 0$ such that

$$N' \doteq \{m' : |\langle m', f_i \rangle - \langle m, f_i \rangle| < \delta, \quad i = 1 \dots n\} \subset N.$$

Let $\delta' < \|m\|^{-1} \min(\epsilon, \delta)$. Since K is compact, there exists a finite number of open sets V_1, \dots, V_r such that $K \subset \cup_{j=1}^r V_j$ and for $j = 1 \dots r$

$$\sup_{x \in V_j} f_i(x) - \inf_{x \in V_j} f_i(x) < \delta', \quad i = 1 \dots n,$$

and

$$\sup_{x \in V_j} f(x) - \inf_{x \in V_j} f(x) < \delta'.$$

By condition (i) of Hypothesis 2, the support of μ equals X . Hence $\mu(V_j) > 0$ for $j = 1 \dots r$. We can find compact sets $K_1, \dots, K_r \subset X$ such that $0 < \mu(K_j) < \infty$ and $K_j \subset V_j$ for $j = 1 \dots r$. This is possible because μ is a Radon measure and each V_j is open. Define subsets U_1, \dots, U_r of X by $U_1 \doteq V_1$ and $U_j \doteq V_j \setminus U_{j-1}$ for $j = 2, \dots, r$. Then $U_j \subset V_j$, $U_j \cap U_{j'} = \emptyset$ for $j \neq j'$ and $K \subset \cup_{j=1}^r U_j$. Let

$$m'(dx) \doteq \sum_{j=1}^r \frac{m(U_j)}{\mu(K_j)} 1_{K_j}(x) \mu(dx).$$

Then

$$\sum_{j=1}^r m(U_j) \inf_{x \in V_j} f_i(x) \leq \int_X f_i(x)m(dx) \leq \sum_{j=1}^r m(U_j) \sup_{x \in V_j} f_i(x)$$

and

$$\sum_{j=1}^r m(U_j) \inf_{x \in V_j} f_i(x) \leq \int_X f_i(x)m(dx) \leq \sum_{j=1}^r m(U_j) \sup_{x \in V_j} f_i(x).$$

Therefore for $i = 1 \dots n$

$$\left| \int_X f_i(x)m(dx) - \int_X f_i(x)m'(dx) \right| \leq \|m\| \delta' < \delta.$$

Thus $m' \in N$. Similarly

$$\left| \int_X f(x)m(dx) - \int_X f(x)m'(dx) \right| < \|m\| \delta' < \epsilon.$$

The proof of the lemma is complete. □

Lemma 3.5 If $m \in E$ satisfies $I(m) < \infty$, then for any neighbourhood N of m and any $\epsilon > 0$ there is a measure $m' \in N$ which is absolutely continuous with respect to μ and which satisfies $|I(m) - I(m')| < \epsilon$.

Proof: If $\gamma = \infty$, then we may choose $m' = m$, and we are done. Now suppose that $\gamma < \infty$. By Lemma 3.3 there exists a measure $\bar{m} \in N$ such that \bar{m}_s has compact support and $|I(\bar{m}) - I(m)| < \frac{1}{2}\epsilon$. Since I is lower semi-continuous, there is a neighbourhood M of \bar{m} such that if $\tilde{m} \in M$, then

$$I(\tilde{m}) > I(\bar{m}) - \frac{1}{2}\epsilon.$$

Thus if $\tilde{m} \in M$, then

$$I(\tilde{m}) > I(m) - \epsilon.$$

Let $L \doteq \{m' \in E : m' + \bar{m}_a \in N \cap M\}$. L is a neighbourhood of \bar{m}_s , and therefore by Lemma 3.4 there exists $m^{(1)} \in L$ such that $m^{(1)}$ is absolutely continuous with respect to μ and satisfies

$$\left| \int_X (\gamma - \xi(x)) m^{(1)}(dx) - \int_X (\gamma - \xi(x)) \bar{m}_s(dx) \right| < \frac{1}{2}\epsilon. \quad (3.6)$$

Let $m^{(2)} \doteq m^{(1)} + m_a$. By definition of L , since $m^{(1)} \in L$, we have $m^{(2)} = m^{(1)} + \bar{m}_a \in N \cap M$ and therefore

$$I(m^{(2)}) \geq I(m) - \epsilon.$$

On the other hand, by formula (2.14), Lemma 2.7 and formula (3.6)

$$\begin{aligned} I(m^{(2)}) &= \sup_{f \in \mathcal{D}} I(m^{(2)}, f) = \sup_{f \in \mathcal{D}} \{ \langle m^{(1)}, f \rangle + I(\bar{m}_a, f) \} \\ &\leq \int_X (\gamma - \xi(x)) m^{(1)}(dx) + I(\bar{m}_a) \\ &< \int_X (\gamma - \xi(x)) \bar{m}_s(dx) + I(\bar{m}_a) + \frac{1}{2}\epsilon \\ &= U(\bar{m}_s) + I(\bar{m}_a) + \frac{1}{2}\epsilon \\ &= I(\bar{m}) + \frac{1}{2}\epsilon \\ &< I(m) + \epsilon. \end{aligned}$$

Thus the measure $m' = m^{(2)} \in N$ is absolutely continuous with respect to μ and satisfies $|I(m) - I(m')| < \epsilon$. This completes the proof of the lemma.

□

Proof of Theorem 1 (b): Part (b) of Theorem 1 follows from part (a) by Lemma 3.5.

□

The proof of Theorem 1 is now complete.

Appendix

Proof of Lemma 1.1

Since π is convex, we have for any two real numbers a and b in $(-\infty, \gamma)$

$$(b - a)\pi'(a) \leq \pi(b) - \pi(a) \leq (b - a)\pi'(b). \quad (A.1)$$

(i) Suppose that $g \in \mathcal{G}$ and choose $\epsilon > 0$ such that $g + \epsilon \in \mathcal{G}$. Putting $a \doteq g(x)$ and $b \doteq g(x) + \epsilon$ in (A.1) we get

$$\epsilon\pi'(g(x)) \leq \pi(g(x) + \epsilon) - \pi(g(x)).$$

Therefore

$$\int_X \pi'(g(x))\mu(dx) \leq \epsilon^{-1} \int_x \{|\pi(g(x) + \epsilon)| + |\pi(g(x))|\} \mu(dx).$$

Thus by Hypothesis 2, condition (ii)

$$\int_X \pi'(g(x))\mu(dx) < \infty.$$

(ii) Let $g \in \mathcal{G}$ and $f \in \mathcal{C}^b(X)$. Choose $\epsilon > 0$ such that $g + \epsilon f \in \mathcal{G}$. Putting $a \doteq g(x)$ and $b \doteq g(x) + \epsilon f(x)$ in (A.1) we get

$$\epsilon f(x)\pi'(g(x)) \leq \pi(g(x) + \epsilon f(x)) - \pi(g(x)).$$

Therefore

$$\int_X f(x)\pi'(g(x))\mu_n(dx) \leq \epsilon^{-1} \int_X \{\pi(g(x) + \epsilon f(x)) - \pi(g(x))\} \mu_n(dx).$$

Thus by Hypothesis 2, condition (iii)

$$\limsup_{n \rightarrow \infty} \int_X f(x)\pi'(g(x))\mu_n(dx) \leq \epsilon^{-1} \int_X \{\pi(g(x) + \epsilon f(x)) - \pi(g(x))\} \mu(dx).$$

Using (A.1) again, we obtain

$$\epsilon^{-1} \int_x \{\pi(g(x) + \epsilon f(x)) - \pi(g(x))\} \mu(dx) \leq \int_X f(x)\pi'(g(x) + \epsilon f(x))\mu(dx).$$

It follows then that

$$\limsup_{n \rightarrow \infty} \int_X f(x)\pi'(g(x))\mu_n(dx) \leq \int_X f(x)\pi'(g(x) + \epsilon f(x))\mu(dx).$$

Using part (i) of this lemma and Lebesgue's Dominated Convergence Theorem, we see that

$$\lim_{\epsilon \rightarrow 0} \int_X f(x) \pi'(g(x) + \epsilon f(x)) \mu(dx) = \int_X f(x) \pi'(g(x)) \mu(dx).$$

Therefore

$$\limsup_{n \rightarrow \infty} \int_X f(x) \pi'(g(x)) \mu_n(dx) \leq \int_X f(x) \pi'(g(x)) \mu(dx).$$

The other inequality

$$\liminf_{n \rightarrow \infty} \int_X f(x) \pi'(g(x)) \mu_n(dx) \geq \int_X f(x) \pi'(g(x)) \mu(dx)$$

follows by a similar argument by putting $b = g(x)$ and $a = g(x) - \epsilon f(x)$ in (A.1).

(iii) Let $\{c_n\}$ be a sequence of real numbers converging to 0. Let $g \in \mathcal{G}$ and $f \in C^b(X)$. We first suppose that $c_n > 0$ for each n . From (A.1) we get for n sufficiently large

$$\begin{aligned} \int_X f(x) \pi'(g(x)) \mu_n(dx) &\leq c_n^{-1} \int_X \{\pi(g(x) + c_n f(x)) - \pi(g(x))\} \mu_n(dx) \\ &\leq \int_X f(x) \pi'(g(x) + c_n f(x)) \mu_n(dx). \end{aligned} \quad (\text{A.2})$$

By part (ii) of this lemma the left hand side of the inequality (A.2) gives immediately

$$\int_X f(x) \pi'(g(x)) \mu(dx) \leq \liminf_{n \rightarrow \infty} c_n^{-1} \int_X \{\pi(g(x) + c_n f(x)) - \pi(g(x))\} \mu_n(dx). \quad (\text{A.3})$$

Given $\epsilon > 0$, then for n large enough so that $c_n \|f\|_\infty < \epsilon$ we get from the right hand side of (A.2)

$$\begin{aligned} c_n^{-1} \int_X \{\pi(g(x) + c_n f(x)) - \pi(g(x))\} \mu_n(dx) &\leq \int_X f_+(x) \pi'(g(x) + \epsilon) \mu_n(dx) - \int_X f_-(x) \pi'(g(x) - \epsilon) \mu_n(dx). \end{aligned}$$

This follows from the fact that π' is an increasing function. Thus by part (ii) of this lemma

$$\begin{aligned} \limsup_{n \rightarrow \infty} c_n^{-1} \int_X \{\pi(g(x) + c_n f(x)) - \pi(g(x))\} \mu_n(dx) &\leq \int_X f_+(x) \pi'(g(x) + \epsilon) \mu(dx) - \int_X f_-(x) \pi'(g(x) - \epsilon) \mu(dx). \end{aligned}$$

By Lebesgue's Dominated Convergence Theorem and part (i) of this lemma we then have

$$\limsup_{n \rightarrow \infty} c_n^{-1} \int_x \{\pi(g(x) + c_n f(x)) - \pi(g(x))\} \mu_n(dx) \leq \int_X f(x) \pi'(g(x)) \mu(dx). \quad (A.4)$$

Combining the inequalities (A.3) and (A.4) we get

$$\lim_{n \rightarrow \infty} c_n^{-1} \int_x \{\pi(g(x) + c_n f(x)) - \pi(g(x))\} \mu_n(dx) = \int_X f(x) \pi'(g(x)) \mu(dx). \quad (A.5)$$

If $c_n < 0$ for each n we can replace c_n by $-c_n$ and f by $-f$ in (A.5) to obtain the same result. This completes the proof of Lemma 1.1.

□

REFERENCES

1. Varadhan, S. R. S.: Asymptotic probabilities and differential equations. Commun. Pure Appl. Math. **19**, 261-286 (1966)
2. Ellis, R. S.: Entropy, Large Deviations and Statistical Mechanics. Berlin, Heidelberg, New York: Springer 1985
3. Deuschel, J. D., Stroock, D. W.: Large Deviations. Boston: Academic Press 1989
4. van den Berg, M., Dorlas, T. C., Lewis, J. T., Pulé, J.V.: A perturbed mean-field model of an interacting boson gas and the large deviation principle. Commun. Math. Phys. **127**, 41-69 (1990)
5. Dorlas, T. C., Lewis, J. T., Pulé, J.V.: Condensation in some perturbed mean-field models of a Bose gas. Helv. Phys. Acta **64**, 1200-1224 (1991)
6. van den Berg, M., Dorlas, T. C., Lewis, J. T., Pulé, J.V.: The pressure in the Huang-Yang-Luttinger model of an interacting boson gas. Commun. Math. Phys. **128**, 231-245 (1990)
7. Bourbaki, N.: Éléments de Mathématiques, Chap. IX: Intégration. Paris: Hermann 1969
8. Rudin, W.: Real and Complex Analysis. New York: McGraw-Hill 1966